STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/550,525_

Source: _PU/o_

Date Processed by STIC: _10/6/05_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A·COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):** U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

Raw Sequence Listing Error Summary

ERROR DETECTED   **SUGGESTED CORRECTION**   SERIAL NUMBER: _10/550,525_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1____Wrapped Nucleics   The number/text at the end of each line "wrapped" down to the next line. This may occur if your file
    Wrapped Aminos   was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will
           prevent "wrapping."

2____Invalid Line Length   The rules require that a line **not exceed** 72 characters in length. This includes white spaces.

3____Misaligned Amino   The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers;
    Numbering   use **space characters**, instead.

4____Non-ASCII   The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please
    ensure your subsequent submission is saved in ASCII text.**

5____Variable Length   Sequence(s)____contain n's or Xaa's representing more than one residue. **Per Sequence Rules,
    each n or Xaa can only represent a single residue.** Please present the **maximum** number of each
    residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____PatentIn 2.0   A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
    "bug"   sequences(s) _____. Normally, PatentIn would automatically generate this section from the
    previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to
    the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for
    Artificial or Unknown sequences.**

7____Skipped Sequences   Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
    (OLD RULES)   (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
    (i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
    (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
    This sequence is intentionally skipped

    Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8____Skipped Sequences   Sequence(s) ____ missing. If **intentional,** please insert the following lines for **each** skipped sequence.
    (NEW RULES)   <210> sequence id number
    <400> sequence id number
    000

9____Use of n's or Xaa's   Use of n's and/or Xaa's have been detected in the Sequence Listing.
    (NEW RULES)   Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
    In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213>   Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or
    Response   scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or
    is Artificial Sequence

11____Use of <220>   Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or
    "Unknown." Please explain source of genetic material in <220> to <223> section.
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12____PatentIn 2.0   Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file,
    "bug"   resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence
    listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____Misuse of n/Xaa   "n" can only represent a single <u>nucleotide</u>; "Xaa" can only represent a single <u>amino acid</u>

AMC – Biotechnology Systems Branch – 09/09/2003

PCT

```
   3 <110> APPLICANT: Shim, Hyunsuk
   4         Liang, Zhongxing
   5         Goodman, Mark
   6         Taichman, Russel
   7         Umbreh, Jay
   9 <120> TITLE OF INVENTION: CXCR4 Antagonists and Methods of Their Use
  11 <130> FILE REFERENCE: 50508-2330 (Emory Ref. No. 02070)
C-->13 <140> CURRENT APPLICATION NUMBER: US/10/550,525
C-->13 <141> CURRENT FILING DATE: 2005-09-22
  13 <150> PRIOR APPLICATION NUMBER: 60/458,217
  14 <151> PRIOR FILING DATE: 2003-03-27
  16 <160> NUMBER OF SEQ ID NOS: 20
  18 <170> SOFTWARE: PatentIn version 3.2
  20 <210> SEQ ID NO: 1
  21 <211> LENGTH: 14
  22 <212> TYPE: PRT
  23 <213> ORGANISM: Artificial
  25 <220> FEATURE:
  26 <223> OTHER INFORMATION: sequence of T140
  29 <220> FEATURE:
  30 <221> NAME/KEY: MISC_FEATURE
  31 <222> LOCATION: (3)..(3)
  32 <223> OTHER INFORMATION: X = Nal
  34 <220> FEATURE:
  35 <221> NAME/KEY: MISC_FEATURE
  36 <222> LOCATION: (8)..(8)
  37 <223> OTHER INFORMATION: X = dLys
  39 <220> FEATURE:
  40 <221> NAME/KEY: MISC_FEATURE
  41 <222> LOCATION: (12)..(12)
  42 <223> OTHER INFORMATION: X = Cit
  44 <400> SEQUENCE: 1
W-->46 Arg Arg Xaa Cys Tyr Arg Lys Xaa Pro Tyr Arg Xaa Cys Arg
  47 1               5                   10
  50 <210> SEQ ID NO: 2
  51 <211> LENGTH: 14
  52 <212> TYPE: PRT
  53 <213> ORGANISM: Artificial
  55 <220> FEATURE:
  56 <223> OTHER INFORMATION: sequence of TN14003
  59 <220> FEATURE:
  60 <221> NAME/KEY: MISC_FEATURE
  61 <222> LOCATION: (3)..(3)
```

Handwritten annotations: pp 1-5 ; Does Not Comply Corrected Diskette Needed ; give source of genetic material (see item 11 on Error Summary Sheet) ; from what is this derived?

```
62 <223> OTHER INFORMATION: X = Nal
64 <220> FEATURE:
65 <221> NAME/KEY: MISC_FEATURE
66 <222> LOCATION: (6)..(6)
67 <223> OTHER INFORMATION: X = Cit
69 <220> FEATURE:
70 <221> NAME/KEY: MISC_FEATURE
71 <222> LOCATION: (8)..(8)
72 <223> OTHER INFORMATION: X = dLys
74 <220> FEATURE:
75 <221> NAME/KEY: MISC_FEATURE
76 <222> LOCATION: (12)..(12)
77 <223> OTHER INFORMATION: X = Cit
79 <400> SEQUENCE: 2
```

W--> `81 Arg Arg Xaa Cys Tyr Xaa Lys Xaa Pro Tyr Arg Xaa Cys Arg`
```
82 1               5                   10
85 <210> SEQ ID NO: 3
86 <211> LENGTH: 18
87 <212> TYPE: PRT
88 <213> ORGANISM: Artificial
90 <220> FEATURE:
91 <223> OTHER INFORMATION: Sequence of T22
94 <220> FEATURE:
95 <221> NAME/KEY: MISC_FEATURE
96 <222> LOCATION: (4)..(4)
97 <223> OTHER INFORMATION: X is Nal which is L-3-(2-napthyl)alanine
99 <400> SEQUENCE: 3
101 Arg Arg Trp Cys Tyr Arg Lys Cys Tyr Lys Gly Tyr Cys Tyr Arg Lys
102 1           5                   10                  15
105 Cys Arg
109 <210> SEQ ID NO: 4
110 <211> LENGTH: 21
111 <212> TYPE: DNA
112 <213> ORGANISM: Artificial
114 <220> FEATURE:
115 <223> OTHER INFORMATION: cDNA sequence segments of CXCR4
117 <400> SEQUENCE: 4
118 aataaaatct tcctgcccac c                                    21
121 <210> SEQ ID NO: 5
122 <211> LENGTH: 21
123 <212> TYPE: DNA
124 <213> ORGANISM: Artificial
126 <220> FEATURE:
127 <223> OTHER INFORMATION: cDNA sequence segments of CXCR4
129 <400> SEQUENCE: 5
130 aaggaagctg ttggctgaaa a                                    21
133 <210> SEQ ID NO: 6
134 <211> LENGTH: 19
135 <212> TYPE: DNA
```

[handwritten annotation: ? "Cys" is at location 4]

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,525   TIME: 11:42:43

Input Set : **A:\050508-1370.txt**
Output Set: **N:\CRF4\10062005\J550525.raw**

```
136 <213> ORGANISM: Artificial
138 <220> FEATURE:
139 <223> OTHER INFORMATION: CXCR4 cDNA target sequence
141 <400> SEQUENCE: 6
142 taactacacc gaggaaatg                                        19
145 <210> SEQ ID NO: 7
146 <211> LENGTH: 19
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial
150 <220> FEATURE:
151 <223> OTHER INFORMATION: CXCR4 cDNA target sequence
153 <400> SEQUENCE: 7
154 tcttcttaac tggcattgt                                        19
157 <210> SEQ ID NO: 8
158 <211> LENGTH: 19
159 <212> TYPE: DNA
160 <213> ORGANISM: Artificial
162 <220> FEATURE:
163 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
165 <400> SEQUENCE: 8
166 tctttgccaa cgtcagtga                                        19
169 <210> SEQ ID NO: 9
170 <211> LENGTH: 19
171 <212> TYPE: DNA
172 <213> ORGANISM: Artificial
174 <220> FEATURE:
175 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
177 <400> SEQUENCE: 9
178 gtttcagcac atcatggtt                                        19
181 <210> SEQ ID NO: 10
182 <211> LENGTH: 19
183 <212> TYPE: DNA
184 <213> ORGANISM: Artificial
186 <220> FEATURE:
187 <223> OTHER INFORMATION: CXCR4 cDNA target sequence
189 <400> SEQUENCE: 10
190 catcatggtt ggccttatc                                        19
193 <210> SEQ ID NO: 11
194 <211> LENGTH: 19
195 <212> TYPE: DNA
196 <213> ORGANISM: Artificial
198 <220> FEATURE:
199 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
201 <400> SEQUENCE: 11
202 tcctgcctgg tattgtcat                                        19
205 <210> SEQ ID NO: 12
206 <211> LENGTH: 19
207 <212> TYPE: DNA
208 <213> ORGANISM: Artificial
```

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,525   TIME: 11:42:43

Input Set : **A:\050508-1370.txt**
Output Set: **N:\CRF4\10062005\J550525.raw**

```
210 <220> FEATURE:
211 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
213 <400> SEQUENCE: 12
214 tcctgtcctg ctattgcat                                    19
217 <210> SEQ ID NO: 13
218 <211> LENGTH: 19
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial
222 <220> FEATURE:
223 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
225 <400> SEQUENCE: 13
226 gcatcgactc cttcatcct                                    19
229 <210> SEQ ID NO: 14
230 <211> LENGTH: 19
231 <212> TYPE: DNA
232 <213> ORGANISM: Artificial
234 <220> FEATURE:
235 <223> OTHER INFORMATION: CXCR4 cDNA target sequences
237 <400> SEQUENCE: 14
238 ggaaagcgag gtggacatt                                    19
241 <210> SEQ ID NO: 15
242 <211> LENGTH: 25
243 <212> TYPE: DNA
244 <213> ORGANISM: Artificial
246 <220> FEATURE:
247 <223> OTHER INFORMATION: siRNA
249 <400> SEQUENCE: 15
250 aauaaaaucu uccugcccac cdtdt                             25
253 <210> SEQ ID NO: 16
254 <211> LENGTH: 25
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial
258 <220> FEATURE:
259 <223> OTHER INFORMATION: siRNA
261 <400> SEQUENCE: 16
262 aaggaagcug uuggcugaaa adtdt                             25
265 <210> SEQ ID NO: 17
266 <211> LENGTH: 20
267 <212> TYPE: DNA
268 <213> ORGANISM: Artificial
270 <220> FEATURE:
271 <223> OTHER INFORMATION: CXCR4 specifc primers
273 <400> SEQUENCE: 17
274 gaaccctgtt tccgtgaaga                                   20
277 <210> SEQ ID NO: 18
278 <211> LENGTH: 20
279 <212> TYPE: DNA
280 <213> ORGANISM: Artificial
282 <220> FEATURE:
```

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: **US/10/550,525**   TIME: 11:42:43

Input Set : **A:\050508-1370.txt**
Output Set: **N:\CRF4\10062005\J550525.raw**

```
283  <223> OTHER INFORMATION: CXCR4-specific primers
285  <400> SEQUENCE: 18
286  cttgtccgtc atgcttctca                                         20
289  <210> SEQ ID NO: 19
290  <211> LENGTH: 20
291  <212> TYPE: DNA
292  <213> ORGANISM: Artificial
294  <220> FEATURE:
295  <223> OTHER INFORMATION: primer
297  <400> SEQUENCE: 19
298  gacaggatgc agaaggagat                                         20
301  <210> SEQ ID NO: 20
302  <211> LENGTH: 20
303  <212> TYPE: DNA
304  <213> ORGANISM: Artificial
306  <220> FEATURE:
307  <223> OTHER INFORMATION: Primer
309  <400> SEQUENCE: 20
310  tgcttgctga tccacatctg                                         20
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,525   TIME: 11:42:44

Input Set : **A:\050508-1370.txt**
Output Set: **N:\CRF4\10062005\J550525.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 3,8,12
Seq#:2; Xaa Pos 3,5,8,12

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20

VERIFICATION SUMMARY   DATE: 10/06/2005
PATENT APPLICATION: **US/10/550,525**   TIME: 11:42:44

Input Set : **A:\050508-1370.txt**
Output Set: **N:\CRF4\10062005\J550525.raw**

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:46 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:81 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0